Structure inference in complex environments improves from childhood to adulthood

Nora C. Harhen[1], Rheza Budiono[1], Catherine A. Hartley[1*], and Aaron M. Bornstein[2,3*]

[1]Department of Psychology, New York University

[2]Department of Cognitive Sciences, University of California, Irvine

[3]Center for the Neurobiology of Learning & Memory, University of California, Irvine

[*]Denotes equal contribution

Study Disclosures

Conflicts of interest: All authors declare no conflicts of interest. Funding: This work was supported by the National Institute of Mental Health (R01 MH126183 to CAH, P50 MH096889 to AMB, PI: TZ Baram, F31 MH134620 to NCH), the Templeton World Charity Foundation (to CAH), the New York University Vulnerable Brain Project (to CAH), and the Department of Defense (NDSEG fellowship to NCH). Ethics: This research received approval from New York University's Institutional Review Board (ID:2021-5210). Preregistration: No aspects of the study were preregistered. Materials: All study materials are publicly available on Github (`https://github.com/noraharhen/Harhen-Hartley-Bornstein-2025-Foraging/tree/main/run_exp`). Data: All primary data are publicly available on Github (`https://github.com/noraharhen/Harhen-Hartley-Bornstein-2025-Foraging/tree/main/data_analysis/data`) Analysis scripts: All analysis scripts are publicly available on Github (`https://github.com/noraharhen/Harhen-Hartley-Bornstein-2025-Foraging/tree/main/data_analysis`).

Abstract

From an early age, children skillfully extract statistical regularities and infer causal relationships. While past work has emphasized children's strong structure learning abilities, it has often relied on tasks that scaffold the learning problem. In these settings, the environmental structure is simple, the transitions between events are repetitive, and participants are explicitly instructed to uncover latent structure. Real-world environments, in contrast, are complex, dynamic, and ambiguous. As children mature into adolescents and then adults, they increasingly face these environments on their own. To compare how learners navigate such conditions across development, we tested 252 participants ages 8 to 25 on a patch-foraging task designed to probe structure learning and planning in a richly-structured environment. Younger participants explored more broadly than adults, leaving patches sooner and sampling a greater number overall. Computational modeling revealed that this difference in exploration stemmed from differences in representation. Younger participants tended to group patches with differing rewards together, while older participants inferred distinct categories based on richness. Despite these representational differences, participants of all ages adjusted their planning based on their internal uncertainty. Our findings suggest that while structure learning continues to develop into young adulthood, uncertainty-sensitive planning emerges early on and supports adaptive behavior even when internal representations of the environment remain imprecise.

Keywords: structure learning, development, foraging, exploration

Research Highlights

- Children and adolescents formed less granular representations of environmental structure than adults.
- Despite these imprecise representations, younger participants showed adult-like sensitivity to uncertainty, planning further ahead when more confident in their internal models of the environment.

Introduction

From an early age, children are adept at carving out structure from streams of sensory information. They extract statistical regularities, infer causal relationships, and even design interventions to test competing hypotheses (Saffran, Aslin, & Newport, 1996; Fiser & Aslin, 2002; Jung, Walther, & Finn, 2021; Bullock, 1985; Leslie, 1982; Gopnik, Sobel, Schulz, & Glymour, 2001; Lapidow & Walker, 2020). These structure-learning abilities allow them to build coherent, causal models of their environments, ones that they can use to guide their predictions, reasoning, and actions. However, most evidence for these abilities comes from tasks whose environments sharply differ from those found in the real world. These tasks are often simple, repetitive, and are sometimes accompanied by explicit instructions to uncover latent structure. But as children grow into adolescents and young adults, such scaffolding becomes less common. They increasingly face complex, ambiguous, and dynamic environments on their own. How do older children and adolescents navigate these settings? Do they spontaneously attempt to infer the environment's true latent structure, even without instruction to do so? How do their internal models compare to those of adults?

In some cases, children have been found to construct more accurate models of the environment than adults. Their broader attention, heightened exploration, and lack of prior knowledge may confer learning advantages (Gualtieri & Finn, 2022). Children are less likely to persist with false beliefs (Liquin & Gopnik, 2022), less prone to schema-driven memory distortions (Blanco & Sloutsky, 2020; Deng & Sloutsky, 2016, 2015), and more likely to discover unconventional causal relationships (Lucas, Bridgers, Griffiths, & Gopnik, 2014). These benefits have largely been observed in children under the age of seven. But, while older children and adolescents are more goal-directed than these younger children, they still continue to explore more than adults (Nussenbaum & Hartley, 2019). Their exploratory tendencies could support the construction of more veridical representations in complex and uncertain environments.

Although children excel at detecting structure, they often struggle to use it. The ability to translate structural knowledge into efficient reasoning and decision making develops gradually across childhood and adolescence (Hartley, Nussenbaum, & Cohen, 2021), perhaps reflecting the prolonged maturation of the prefrontal cortex and anterior hippocampus (Davidow, Insel, & Somerville, 2018; Schlichting & Preston, 2015; DeMaster, Pathman, Lee, & Ghetti, 2014). Consistent with this protracted developmental trajectory, explicit judgments of structure continue to improve between the ages of five and twelve (Arciuli & Simpson, 2011; Shufaniya & Arnon, 2018; Raviv & Arnon, 2018) while implicit indications of sensitivity to structure emerge much earlier (Forest, Schlichting, Duncan, & Finn, 2023). While children and adolescents may be capable of inferring complex structure, they may also find it more costly to deploy, due to still-developing cognitive control abilities and less explicit knowledge representations. As a result, they could default to simpler and less veridical representations that are easier to manage during decision making.

We tested these two possibilities using a patch-foraging task with a complex and uninstructed underlying structure. In general, patch-foraging tasks are thought to capture ecologically relevant learning and choice behaviors (Mobbs, Trimmer, Blumstein, & Dayan, 2018). This variant of the task was designed to reveal how learners represent the environment's structure (Harhen & Bornstein, 2023). Participants, ages 8 to 25, made repeated decisions about whether to keep harvesting a depleting patch of resources or leave in search of a potentially better one. Unbeknownst to them, the environment contained three types of patches—poor, neutral, and rich—each associated with a different distribution over depletion rates. Participants were not told about these patch types, nor were they explicitly instructed to learn their structure. They were only tasked with gathering as many resources as possible within a fixed time window. To infer participants' internal representations, we fit their choices to a Bayesian latent inference model (Harhen & Bornstein, 2023). Because different representations produce distinct stay-leave patterns, the model allows us to identify the most likely representation each participant used to guide their decisions. Because participants can never be certain of how accurate their representation is, they might also incorporate this uncertainty into their decision making (Jiang, Kulesza, Singh, & Lewis, 2015). Our model also allows us to measure the extent to which participants leveraged their representational uncertainty to inform their choices.

Methods

Participants

Our final sample consisted of 252 participants, aged 8 to 25 (M = 17.11 years, SD = 5.29, 128 females, 124 males). This sample included 70 children (8.08 - 12.94 years; M = 10.49, 36 females), 68 adolescents (13.07 - 17.94 years; M = 15.47, 35 females), and 114 adults (18 - 25.83 years; M = 22.14, 57 females). This sample size exceeds those of many prior developmental studies investigating value-guided learning and decision making (Cohen, Nussenbaum, Dorfman, Gershman, & Hartley, 2020; Nussenbaum, Scheuplein, Phaneuf, Evans, & Hartley, 2020; Nussenbaum et al., 2023). We selected this sample size to achieve at least 80% power to detect a small effect size (reflecting individual differences in structure learning) taken from the original adult study (ρ=0.16, (Harhen & Bornstein, 2023)) with α=0.05.

All participants reported normal or corrected-to-normal vision and no history of psychiatric or learning disorders. Based on self- or parent-report, 42% of participants were White, 31% were Asian, 13.1% were mixed race, 12.7% were Black, less than 1% were Pacific Islander or Native Hawaiian, and less than 1% were Native American. Additionally, 11.8% of the sample identified as Hispanic. Participants' annual household incomes ranged from less than \$20,000 to over \$500,000. The distribution was as follows: 9.3% earned less than \$20,000; 7.7% earned \$20,000–39,999; 15.9% earned \$40,000–59,999; 11.4% earned \$60,000–79,999; 11.8% earned \$80,000–99,999; 18.3% earned \$100,000–199,999; 4.9% earned \$200,000–299,999; and fewer than 1% each reported incomes of \$300,000–399,999, \$400,000–499,999, or above \$500,000.

An additional 45 participants completed the study but were excluded based on the following criteria: difficulty understanding task instructions (failing the instruction comprehension check more than twice, n=4), unusually quick responses (mean reaction time < 200 ms, n=12), and extreme strategies (n = 14 for mean planet residence time ±2 *SD* from group mean; n=4 for fully depleting gem mines on more than 75% of visited planets; n=11 for leaving more than 75% of visited planets immediately after the initial dig). Participants received a \$10 Amazon gift card for completing the study and had the opportunity to earn an additional performance-based bonus of up to \$2.

We recruited participants through the Hartley lab's database, for which we solicited sign-ups via Facebook and Instagram ads, local science fairs and events, and fliers on New York University's campus. Lab researchers verified each participant's age and identity prior to their participation in the online study.

Task

Participants completed a child-friendly variant of a patch foraging task previously used to examine structure learning in adults (Harhen & Bornstein, 2023). Modifications were made to accommodate younger participants. These modifications included shortening the length of the task (from five blocks to four), enhancing the clarity of the instructions (see supplemental materials S1.1), augmenting an instruction comprehension quiz (S1.1), and increasing the maximum decision time (from two seconds to three).

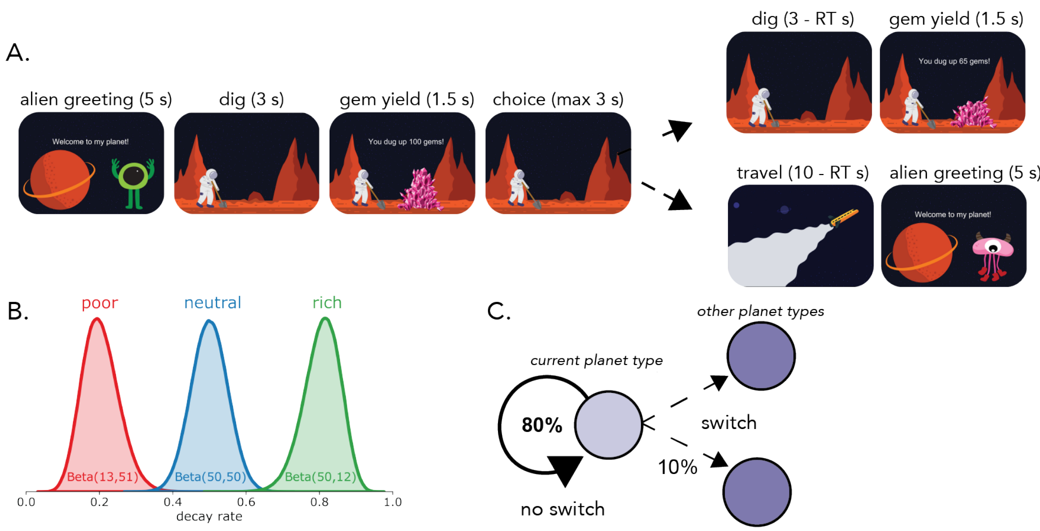

Figure 1: **A. Task design.** Participants traveled to various planets to dig for space gems. On each trial, they decided between continuing to dig on the current planet or traveling to a new one. Both options had their costs: digging depleted the mine, progressively reducing its gem yield, while traveling took a substantial amount of time. **B. Environment structure.** Planets belonged to one of three types–poor, neutral, or rich–differing in how quickly they depleted as mined. Each type was characterized by a distinct distribution over decay rates. **C. Environment dynamics.** Planet richness was correlated in time. A new planet had an 80% probability of being the same type as the previous planet ("no switch") and a 20% probability of transitioning ("switch") to a different type.

In the task, participants acted as miners collecting "space gems" across various planets (Fig. 1). Bonus payments were tied directly to the total amount of gems collected. Upon landing on a planet, participants dug once and received a gem yield sampled from a Normal distribution ($\mu = 100$, $\sigma = 5$). Subsequent trials on the planet involved deciding whether to stay and continue digging, despite increasingly diminishing yields (decreasing according to a planet-type-specific function, described below), or leave for a new planet, incurring a substantial time cost.

If participants chose to stay, a short animation of their avatar digging was shown (3 sec minus the reaction time for that trial), followed by the gem yield (1.5 sec). If they chose to leave, a longer animation of a rocket ship played (10 seconds minus reaction time), followed by an alien welcoming them to the new planet (5.5 sec). Animation durations were linked to the previous trial's reaction time to ensure decision speed did not affect the overall reward rate. Participants had three seconds to respond. If they failed to do so, a red X appeared along with a prompt encouraging them to respond faster. After this, they could re-attempt their choice. In total, the task consisted of four blocks, each lasting six minutes.

The task environment featured three planet types. Rich planets depleted the most slowly (gem yield decreasing with a decay constant sampled from a Beta distribution with parameters α=50, β=12, M=0.8, SD=0.05), poor the most rapidly (α=13, β=51, M=0.2, SD=0.05), and neutral planets at an intermediate rate (α=50, β=50, M=0.5, SD=0.05). On each trial, the depletion rate was newly sampled from the planet's respective distribution. To mimic the structure of natural environments, planet richness was temporally correlated. There was an 80% probability that the next planet would be the same type as the previous one and a 20% chance it would switch to a different type. Importantly, these differences between planets were not explicitly communicated to participants, requiring them to infer this information based on the sequence of observed rewards.

We designed the task to measure participants' representational biases independent of task demands. To achieve this, we structured the sequence of planets such that using a more simple representation, one that grouped all planets together, achieved comparable overall rewards to using a more complex representation, one distinguishing between different planet types. Simulations confirmed that the two strategies yielded similar outcomes on average (see supplemental materials S1.2).

Analysis approach

Mixed effects models We used the "lme4" package for R (Bates, Kliegl, Vasishth, & Baayen, 2018) to fit mixed-effects models to our data. Except where noted, models included participant-level random intercepts and random slopes across within-participant fixed effects. To minimize Type I error, we initially specified the maximal model (Barr, Levy, Scheepers, & Tily, 2013). If the model failed to converge, we iteratively simplified the model by first removing interactions between random slopes, followed by random slopes themselves, until convergence was achieved. We used the 'bobyqa' optimizer and set the number of model iterations to 10,000. Continuous variables–age, planet number, and reaction time–were z-scored prior to their inclusion. Age was z-scored across participants while planet number and reaction times were z-scored within. Reaction times were log-transformed before z-scoring.

Marginal Value Theorem To assess the extent to which individuals over- or under-harvested, we compared their planet (patch) resident times to predictions from the Marginal Value Theorem (Charnov, 1976). According to MVT, an optimal agent decides to stay or leave by comparing the immediate expected returns from staying, (V_{stay}), to the opportunity cost of digging on the current planet, (V_{leave}).

V_{stay} is the reward expected from the next dig, the previous reward multiplied by the predicted depletion rate. An MVT-optimal forager accurately identifies the planet's type and uses the true mean of its depletion distribution for their prediction.

$$V_{stay} = r_t * \hat{d} \tag{1}$$

$$\hat{d} = \begin{cases} 0.2 & \text{if planet is poor} \\ 0.5 & \text{if planet is neutral} \\ 0.8 & \text{if planet is rich} \end{cases}$$

Where r_t is the reward received on the last dig, and $\hat{d}$ is the predicted depletion.

They estimate V_{leave}, the expected reward from digging on an alternative planet using the global reward rate, the total rewards received (r_{total}) divided by the total time spent foraging (t_{total}). Note that under this definition V_{leave} will change with each decision to harvest or leave a planet. By multiplying the global reward rate by the time required to dig (t_{dig}), V_{leave} reflects the opportunity cost of digging on the current planet over an unknown alternative planet.

$$V_{leave} = \frac{r_{total}}{t_{total}} * t_{dig} \tag{2}$$

The forager compares these values and chooses greedily, always selecting the higher-valued action.

Structure learning and uncertainty-adaptive planning model Our model relaxes MVT's assumption of perfect knowledge, introducing two novel computations into the forager's decision-making process: structure learning and uncertainty-adaptive planning.

Foragers must navigate the environment without knowing the true number of planet types, the classification of individual planets, or the decay rate distribution defining each planet type. To model how foragers infer this information, we use the Chinese Restaurant Process (CRP; (Aldous, 1985)). While developed in statistics, the Chinese Restaurant Process has been used as a formal account of numerous psychological processes including category learning (Griffiths, Navarro, & Sanborn, 2006; Perfors & Tenenbaum, 2009; Kemp, Perfors, & Tenenbaum, 2007), state space inference (Gershman, Blei, & Niv, 2010), and the organization of experiences in memory (Shin & DuBrow, 2021). The CRP's prior is defined around two principles: first, the probability of a planet belonging to an existing type increases with the number of planets already assigned to that type, and second, there remains some probability, proportional to the parameter α, of discovering an entirely new type. This model allows the complexity of foragers' representations to grow as they accumulate experience.

$$P(k) = \begin{cases} \frac{n_k}{N+\alpha} & \text{if k is old} \\ \\ \frac{\alpha}{N+\alpha} & \text{if k is new} \end{cases}$$

Where n_k is the number of planets assigned to type k, α is a clustering parameter, and N is the total number of planets encountered.

After observing a depletion on a planet, the forager computes the posterior probability of the planet being a type as:

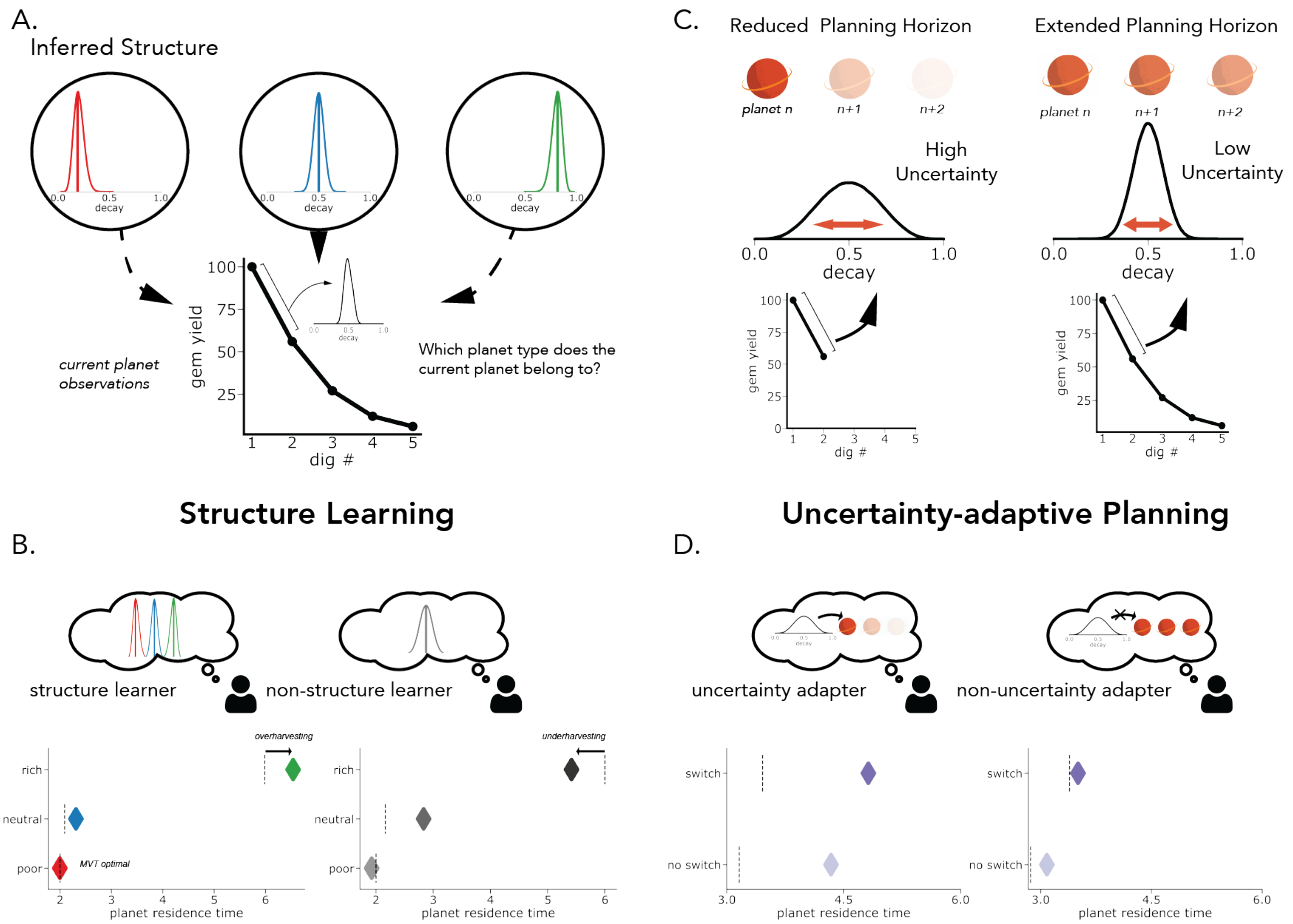

Figure 2: **A. Structure learning computation.** The forager makes two simultaneous inferences: (1) identifying the current planet's type based on observed rewards and (2) determining the total number of planet types in the environment. We model these two inferences as a Chinese Restaurant Process. **B. Structure learning predictions.** The forager's inference of planet types is governed by the parameter, α. The model predicts distinct patterns of over- and under-harvesting depending on the forager's representation of the environment and the number of planet types they consider. The structure learner's behavior is simulated with $\alpha = 0.2$ and the non-structure learner's behavior is simulated with $\alpha = 0$. The markers show model-simulated planet residence times (PRTs), while dotted lines indicate the Marginal Value Theorem (MVT)-optimal PRTs for reference. **C. Uncertainty adaptive planning computation.** Foragers adjust their planning horizon based on their uncertainty about the current planet's type. Less uncertainty encourages planning further into the future, while more uncertainty discourages planning. We modulate the planning horizon through adjusting the extent future rewards are discounted. **D. Uncertainty adaptive planning predictions.** Foragers whose planning incorporates their uncertainty should overharvest more when the planet type switches, because these switches are rare and uncued and so should, on balance, increase local uncertainty. In contrast, foragers who do not incorporate uncertainty into their planning (or who do not detect the change) should overharvest to a similar extent regardless of whether the planet type has changed. The uncertainty adapter's behavior is simulated with $\gamma_{coef} = -0.3$ and the non-uncertainty adapter's behavior is simulated with $\gamma_{coef} = 0$.

$$P(k|D) = \frac{P(D|k)P(k)}{\sum_{j=1}^{J} P(D|j)P(j)} \tag{3}$$

Where J is the number of clusters created up until the current planet, D is a vector of all the depletions observed on the current planet, and all probabilities are conditioned on prior type assignments of planets, $p_{1:N}$

Because computing the exact posterior probability of planet type assignments is computationally intractable, we approximate it using a particle filter. Each particle maintains a hypothetical set of assignments and is weighted based on how well it explains the observed data. When a forager leaves a planet, the particles are resampled. A new pool of particles is generated by sampling with replacement from the previous pool, with the likelihood of a particle being selected proportional to its weight. This process favors particles that better explain the data, increasing their probability of persisting in subsequent iterations. We use a moderate number of particles, 200, as this amount allows for psychological plausibility while still approximating the posterior well (Griffiths et al., 2006).

To predict the next dig's yield, we use a three-step Monte Carlo sampling procedure. First, a particle is selected based on its weight. Second, a planet type is sampled from the selected particle's posterior. Third, a decay rate is drawn from the selected planet type's decay rate distribution. This process is repeated 1,000 times, and the decay rates are averaged over to produce the final prediction.

Each decay rate distribution is initialized to a Gaussian with μ=0.5 and σ=0.5. While the true decay rates follow a Beta distribution, the model assumes normally distributed observations to allow for analytic updates using a Normal-Gamma prior.

Unlike MVT-optimal foragers, who have perfect knowledge, foragers acting according to our model make decisions under epistemic uncertainty. They can never be certain that they have an accurate model of the environment. Theoretical work in reinforcement learning suggests that under uncertainty, reducing how far out into the future the individual plans, the planning horizon, can improve performance (Jiang et al., 2015). We implement this concept in our model with an adaptive factor for discounting future rewards, $\gamma_{effective}$. Note that while the planning horizon technically defines how far into the future the consequences of a decision are evaluated, discounting future consequences produces the same behavior in our task. We define $\gamma_{effective}$ to be a function of an individual's baseline discounting rate,(γ_{base}), their uncertainty over the current planet's type, (U), and the degree to which uncertainty influences their planning,(γ_{coef}). We compute uncertainty (U) as the entropy of the Multinomial distribution over planet types, using the same three-step Monte Carlo sampling procedure described above.

$$\gamma_{effective} = \frac{1}{1 + e^{-(\gamma_{base} + \gamma_{coef} * U)}} \tag{4}$$

Thus, the lower γ_{coef} is the more heavily the forager discounts future rewards. γ_{coef} provides us with an additional measure of the extent a forager uses their structural knowledge, or lack thereof, to adjust their decision making. We included this because of the known developmental dissociations between learning and using structural knowledge.

To model action selection, we use a softmax function, incorporating a lapse rate (ϵ) to account for occasional inattention:

$$p_{stay} = (1 - \epsilon)\frac{1}{1 + e^{-\beta(v_{stay} - v_{leave})}} + \frac{\epsilon}{2} \tag{5}$$

With β being the inverse temperature and ϵ being the lapse rate. Including a lapse rate helps to separate systematic behavior from noise, thus reducing the misestimation of our main parameters of interest.

We compared two versions of the model: one with α fixed at 0 and another with α fixed at 0.2. At α=0, the forager assumes that all planets belong to a single type. α=0.2 was the value that, in simulation, produced the most veridical representation across a range of the model's other parameters (see supplemental materials S1.2). We treated γ_{base}, γ_{coef}, β, and ϵ as free parameters.

Model fitting We fit participants' data on a choice-by-choice basis. Free parameters and their bounds are detailed in the supplementary materials (Table S1). To identify the parameter values that minimized the negative log likelihood of participants' choices, we used Bayesian Adaptive Direct Search (BADS, Acerbi & Ji, 2017), an optimization algorithm suited for stochastic and computationally expensive functions.

To increase the likelihood of finding the global minimum, we initialized the optimization with different starting points generated from a Sobol sequence. Sobol sequences are quasi-random and have been shown to be more effective than grid or random search (Bergstra & Bengio, 2012), while offering greater computational efficiency than Latin hypercube sampling (Renardy, Joslyn, Millar, & Kirschner, 2021). Starting points were generated until the convergence criteria were met, defined as five consecutive iterations without improvement to the overall minimum.

The final parameter values were those that yielded the lowest negative log likelihood across all starting points. Parameter recoverability analyses for both models are included in the supplementary materials (S1.3).

The models' likelihoods are stochastic due to the approximation of the posterior distribution over planet type assignments. To address this noise, we repeated the cluster assignment process 1,000 times. We computed the log likelihood of participants' choices for each of these repetitions and marginalized over them. We then negated this value to obtain the input to the optimization algorithm.

Model comparison We assessed model fit using Akaike Information Criteria (AIC), which penalizes for model complexity. For age-group-level comparisons, we used protected exceedance probabilities (PXP). PXPs estimate the likelihood that a given model is the most frequent best-fitting model within a group while accounting for chance differences in model frequencies. Model recoverability analyses are included in the supplementary materials (S1.3).

Data and code availability statement All data and code are available at `https://github.com/noraharhen/Harhen-Budiono-Hartley-Bornstein-2025-Foraging`.

Results

Computational model-agnostic

Overharvesting increases with age We first examined how overharvesting varied with age. Using a mixed-effects linear regression model, we measured the extent to which participants' planet residence times deviated from Marginal Value Theorem (MVT) optimality. On average, participants stayed longer than the optimal residence time prescribed by MVT (β_0=0.81, SE=0.11, t(248.49)=7.47, p < .001). In other words, they systematically overharvested. Younger participants, however, left planets earlier, aligning more closely with MVT (β_{age}=0.22, SE=0.11, t(248.33)=2.01, p = .045).

Use of structure knowledge strengthens with age We next tested two key predictions of our computational model. First, the model predicts that all foragers, regardless of their representation, will overharvest on neutral planets and less so on poor planets. In contrast, behavior on rich planets should depend on the forager's representation of the environment (Fig. 2B). Foragers who distinguish planets by their type should overharvest the most on neutral planets, moderately on rich planets, and least on poor ones. Those who do not differentiate between types should underharvest on rich planets. By treating all planets as the same, they misestimate how quickly the rich planets will deplete and leave too soon. Second, the model predicts that with increasing experience, learners should form more accurate representations and thus behave more in line with MVT.

To test these predictions, we examined how overharvesting varied with planet type, task experience (indexed by planet number), age, and their interactions. Results from the linear regression model supported both predictions. First, participants overharvested the most on neutral planets, the least on poor planets, and to a moderate extent on rich planets (Fig. 3; intercept: β=1.30, SE=0.081, $t(245.63)$=15.96, $p < .001$; poor planet: β=-0.63, SE=0.052, $t(307.32)$=-12.05, $p < .001$; rich planet: β=-0.42, SE=0.13, $t(245.17)$=-3.15, $p = .0018$). This pattern suggests that, as a group, participants distinguished between planet types. Second, overharvesting decreased with task experience, especially on rich planets (planet number: β=-0.24, SE=0.05, $t(238.88)$=-4.85, $p < .001$; planet number x poor planet interaction: β=0.067, SE=0.047, $t(346.64)$=1.43, $p = .15$; planet number x rich planet interaction: β=-0.26, SE=0.060, $t(229.28)$=-4.29, $p < .001$). As learners accumulated experience, their representations of the environment became more accurate.

These effects varied with age. Older participants overharvested more than younger participants specifically on rich planets (age x rich planet interaction: β=0.36, SE=0.13, $t(245.01)$=2.68, $p = .0078$; age: β=0.059, SE=0.082, $t(245.48)$=0.72, $p = .47$; age x poor planet interaction: β=-0.045, SE=0.052, $t(303.047)$=-0.86, $p = .39$). Older participants' behavior suggested enhanced discrimination between planet types. Overall, these results support our model's predictions and suggests that the use of complex structural knowledge improves with age .

Implicit structure knowledge is present across all ages To measure structure learning more implicitly, we analyzed reaction times following switches in planet type. We focused on participants' second decisions on each planet, as the first depletion provides the initial cue to a planet's type. If participants inferred that planets of the same type cluster together in time, then they should be surprised when a switch occurs, and this should be reflected in slower reaction times.

To test this prediction, we examined how reaction times varied with switches in planet type, planet number, age, and their interactions using a mixed-effects model, including participant-level random intercepts and slopes for the planet switch regressor. Consistent with our prediction, participants responded more slowly following a switch in planet type (switch point: β=0.049, SE=0.023, t(255.5)=2.09, p = .037). This provides additional evidence that participants were sensitive to the environment's structure and dynamics. Reaction times decreased across the task overall (planet number: β=-0.049, SE=0.012, t(8551)=-4.072, $p < .001$), but switch-related slowing did not diminish with experience (switch x planet number interaction: β=0.014, SE=0.024, t(8711)=0.60, p = .55). Importantly, age did not moderate the effect of switches on reaction times (age x switch point interaction: β=0.0088, SE=0.023, t(256.3)=0.38, p = .71). Younger participants slowed down after a switch just like adults, suggesting that an implicit awareness that planets differed was present across our full age range. These results point to a dissociation between detecting structure and the ability to use that structure to guide decision making. While participants of all ages were sensitive to the environment's dynamics, only older participants leveraged this knowledge to inform their choices.

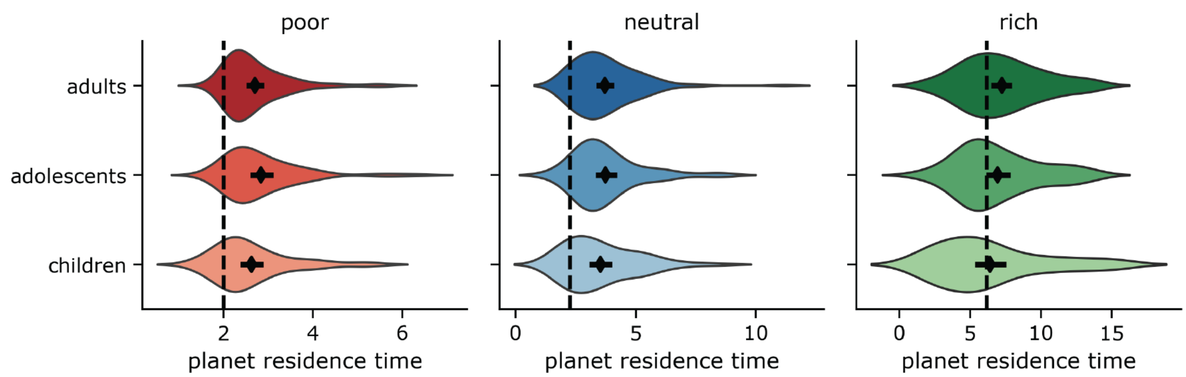

Figure 3: **Model-free signature of structure learning.** Planet residence times (PRT) relative to Marginal Value Theorem (MVT)-optimal PRTs across planet types and age groups. Violin plots show the distribution of PRTs by age group, with markers indicating the means and error bars denoting 95% confidence intervals. Dotted lines provide the MVT-optimal PRT for reference. Overall, participants overharvested across all planet types. Age-related differences were the most pronounced on rich planets. Older participants overharvested on these planets, while many child participants underharvested. Greater variance across child participants on rich planets resulted in a group average PRT that was near MVT-optimal. Our model predictions (Fig 2) suggest that these patterns may reflect age-related differences in the inferred structure of the environment.

Uncertainty adaptive planning emerges early in development We next tested whether participants adjusted their planning horizons based on uncertainty about a planet's type, the key feature of our model's uncertainty-adaptive planning computation. Foragers using this strategy should overharvest more following a switch in planet type, when uncertainty is highest (Fig 4). To test this, we examined how overharvesting varied with switches, task experience, age and their interactions. The model included participant-level random intercepts and slopes for planet number. As predicted, participants overharvested more following switches in planet type (switch point: β=0.31, SE=0.039, $t(8410)$=7.84, $p < .001$). This effect marginally decreased with experience (switch point $\times$ planet number: β=-0.078, SE=0.042, $t(8414)$=-1.85, $p = .065$), suggesting that as participants gained familiarity with the environment, switches were less uncertainty-inducing. Early in the task, switch-related overharvesting did not significantly differ with age (age $\times$ switch point: β=-0.0094, SE=0.039, $t(8410)$=-0.24, $p = .81$). However, older participants showed a greater reduction in switch-related overharvesting with experience (age $\times$ switch point $\times$ planet number: β=-0.11, SE=0.042, $t(8412)$=-2.65, $p = .0082$). These findings suggest that uncertainty-adaptive planning emerges early in development. But, older participants more effectively integrate their knowledge of the environment into their decisions over time.

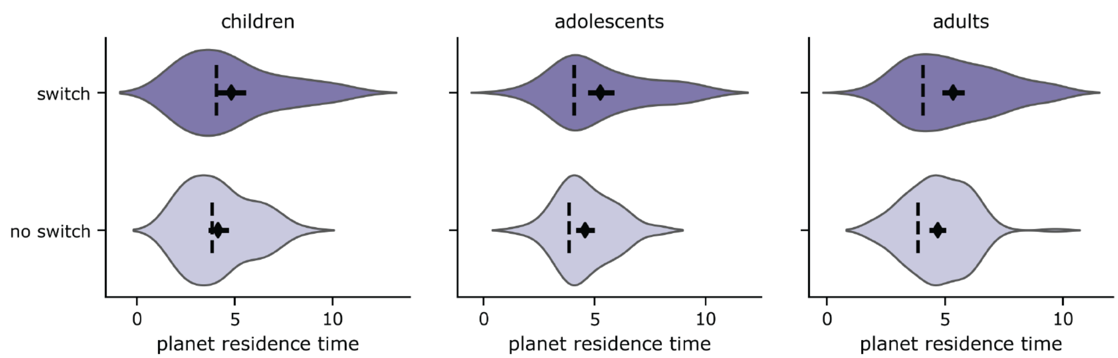

Figure 4: **Model-free signature of uncertainty-adaptive planning.** Planet residence times (PRT) relative to Marginal Value Theorem (MVT)-optimal PRTs across planet types and age groups. Violin plots show the distribution of PRTs by age group, with markers indicating the means and error bars denoting 95% confidence intervals. Dotted lines provide the MVT-optimal PRT for reference. Participants across all age groups adjusted their behavior in response to uncertainty, overharvesting more following rare switches in planet type when uncertainty increased. This behavior is consistent with a reduction in planning horizon.

Computational model-based results

Our central question was whether structure learning differs with age. To address this, we fit two models to participants' choices. In one model, α was fixed at 0, corresponding to an undifferentiated representation of the environment, and in the other, α was fixed at 0.2, the value that most consistently produced a veridical representation in simulations (S1.2). Accordingly, we refer to this model as α^* from here on. To compare these models, we calculated the protected exceedance probabilities (PXP) within each age group (Fig. 5A). Adults' behavior was better captured by the α^* model (PXP = 0.88). They learned the environment's structure and used it to guide their decisions. In contrast, children's choices were better captured by the α=0 model (PXP = 0.83). They did not learn to differentiate between the different planet types. Finally, the most frequent, best-fitting model across adolescents was less clear (α^* PXP = 0.66, α=0 PXP= 0.34). To examine these age-related differences more continuously, we computed the difference in AIC scores between the two models. A positive value indicates the participant's choices are better fit by the α^* model. We found that the difference in AIC scores became increasingly positive with age (Fig 5B, Spearman's ρ = .22, $p < .001$). This recapitulates what we found in our model-agnostic analyses. The ability to infer and use structural knowledge strengthens across development.

Under the α^* model, we found that the baseline discounting factor, the uncertainty-adaptive discounting parameter, the softmax temperature, and the lapse rate did not significantly vary with age (γ_{base}: ρ=-0.049,$p = .44$; γ_{coef}: ρ=-0.068, $p = .28$; β: ρ=0.056, $p = .37$; ϵ: ρ=-0.092, $p = .15$). Fit parameters from the α=0 model yielded the same results (γ_{base} :ρ=-0.041, $p = .52$; β: ρ=0.0095, $p = .88$; ϵ: ρ=-0.087, $p = .17$).

Discussion

Here, we investigated how the ability to learn and leverage latent structure in complex environments changes with age. In our patch-foraging task, younger participants tended to leave planets earlier than adults. This was particularly the case on the richest planets. Our computational model suggests that this difference stems from how participants represented the environment. Younger participants behaved as if they categorized all planets into a single type, while adults appeared to make fine-grained distinctions between planets. Nevertheless, participants of all ages adapted their planning in response to uncertainty, staying longer following unexpected switches in planet type. Taken together, our findings indicate that although the ability to construct and use environmental representations continues to improve into adulthood, even young learners are capable of deploying sophisticated strategies sensitive to the inferred dynamics of their environments.

Children's choices suggest that they represented all planets similarly. However, their response times tell a different story. They slowed down following unexpected changes in planet type. This would suggest then that, on some level, they registered the differences between planets. Why, then, did they not use this knowledge to guide their

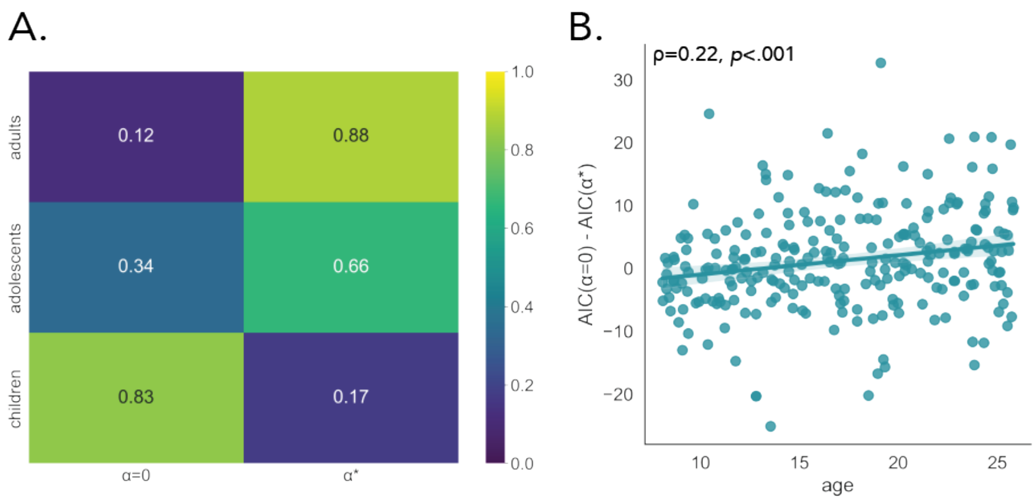

Figure 5: **Model-based results. A.** Protected exceedance probabilities (PXPs) reveal differences in the best-fitting model within age groups. The α^* model best captured adults' choices, while the α=0 model better captured children's choices. But, for adolescents, the best-fitting model was less clear. The α^* had a slightly higher PXP than the α=0 model. **B.** We also examined continuous, rather than group-level, age-related differences in model fit. We computed the difference in Akaike Information Criterias (AICs) for each participant and correlated the difference scores with age. A positive difference indicates that the α^* model better captured a participant's choices. We found that the degree to which the α^* model provided a better fit increased with age (Spearman's ρ=0.22, $p < .001$).

decisions? One possibility is that their structural knowledge remains implicit rather than explicit. Young children often show sensitivity to statistical structure before they can explicitly report or strategically apply this knowledge (Karmiloff-Smith, 1995). Another possibility is that *using* a complex representation is too cognitively demanding. Prior work has found that children and adolescents can acquire structural knowledge but fail to apply it in their planning and reasoning (Nussenbaum et al., 2020; Potter, Bryce, & Hartley, 2017; Decker, Otto, Daw, & Hartley, 2016; Cohen et al., 2020; Schlichting, Guarino, Schapiro, Turk-Browne, & Preston, 2017). This dissociation may reflect the ongoing development of cognitive abilities that underpin the use of mental representations. These include working memory, future simulation, and episodic memory (Hartley et al., 2021), all of which continue to mature into young adulthood (Coughlin, Robins, & Ghetti, 2019; Keresztes, Ngo, Lindenberger, Werkle-Bergner, & Newcombe, 2018; Bunge & Wright, 2007). From this perspective, learning complex representations but using simpler ones may serve an adaptive purpose. It allows children and adolescents to make sufficiently good decisions with respect to the constraints of their cognitive abilities. As these abilities improve, they can begin to take advantage of more complex, sophisticated models.

Alternatively, children may use simpler representations not because of their cognitive constraints, but because it is the rational choice in this context. We designed our task to ensure that both simple and complex representations yielded equal amounts of reward on average. Given this, the simpler representation is the more efficient option, achieving the same outcome but at a lower computational cost. A better question, then, is why adults choose to use the costlier representation. Notably, this preference for costlier representations is not unique to our task. Prior work has found that adults prefer to plan over an internal model of the environment, even when a simpler, model-free strategy would suffice and yield the same payoffs. In contrast, children do opt for the simpler strategy (Nussenbaum et al., 2020; Potter et al., 2017; Decker et al., 2016). One possibility is that adults generalize strategies from their real-world experiences, where uncovering the veridical latent structure and planning ahead is often beneficial. Thus, they might default to these strategies, not because the task explicitly incentivizes it, but rather because their prior experience has taught them it is advantageous.

Notably, even the youngest participants engaged in uncertainty-adaptive planning, remaining longer on planets after unexpected switches in type. While children generally engage in planning less often than adults (Nussenbaum et al., 2020; Potter et al., 2017; Decker et al., 2016), they may engage specifically in uncertainty-adaptive planning because it conserves cognitive resources. Planning deeply when uncertain is ultimately inefficient because predictions are less likely to be accurate. By strategically planning only when certain, children make efficient use of their limited working memory capacity and future simulation abilities. Importantly, this strategy primarily relies on reactive rather than proactive control: cognitive control is only engaged when necessary rather than being continuously deployed in anticipation. This suits children's response tendencies, as children tend to favor reactive control over proactive control (Chevalier, Dauvier, & Blaye, 2018; Chevalier, Meaney, Traut, & Munakata, 2020; Niebaum, Chevalier, Guild, & Munakata, 2021). Additionally, this strategy depends on tracking internal uncertainty, an ability that has been observed in children far younger than those in our sample (Baer & Kidd, 2022; Lapidow, Killeen, & Walker, 2022; Schulz, Wu, Ruggeri, & Meder, 2019; de Eccher, Mundry, & Mani, 2024). Collectively, these findings suggest that our younger participants have the requisite metacognitive abilities to use uncertainty-adaptive planning strategies effectively.

Prior work has found that, in some instances, children are better at discovering latent structure than adults. This has been attributed, in part, to their heightened and broader exploration. While children did explore more broadly in our task, leaving planets sooner, they did not develop more veridical representations. One possibility is that broader exploration may *not* be well-suited to discovering structure in certain environments. Under standard definitions, exploratory actions are ones that provide information at the cost of immediate reward (Schulz & Gershman, 2019). Staying longer under uncertainty fits this definition. It allows learners to gather more observations (i.e., depletions) to refine their estimates of the local patch environment. At the same time, it causes them to deviate from the MVT-optimal behavior that would maximize their immediate reward rate. In patch-foraging tasks, leaving sooner improves estimates of the global reward rate. Staying longer, in contrast, refines knowledge about the local depletion rate. In this sense, both behaviors–leaving sooner and staying longer–serve an exploratory function, but do so in diverging ways. Leaving sooner promotes broad exploration while staying longer promotes deep exploration (Moreno-Bote, Ramírez-Ruiz, Drugowitsch, & Hayden, 2020). Thus, counter to conventional conceptualizations, we propose that over-exploitation may actually be a form of exploration and one that is better-suited to patch-foraging environments. That these distinct forms of exploration exhibit changes over development raises the question of how each strategy may facilitate the acquisition of different forms of structural knowledge.

Our findings contribute new insights into how structure learning and its use in complex environments change across development. Through leveraging a Bayesian structure learning model, we were able to infer participants' representations of the foraging environment. Significantly, children's broader exploration did not confer a learning advantage. Instead, adults' overexploitation proved to be beneficial for uncovering the environment's true structure. The current study paves the way for future work examining how age-related differences in information-seeking interact with the structure of different environments to produce representational differences across development.

References

Acerbi, L., & Ji, W. (2017). Practical bayesian optimization for model fitting with bayesian adaptive direct search. *Adv. Neural Inf. Process. Syst.*, 1836–1846.

Aldous, D. J. (1985). Exchangeability and related topics. In D. J. Aldous, I. A. Ibragimov, & J. Jacod (Eds.), *École d'été de probabilités de Saint-Flour XIII — 1983* (pp. 1–198). Springer Berlin Heidelberg.

Arciuli, J., & Simpson, I. C. (2011, May). Statistical learning in typically developing children: the role of age and speed of stimulus presentation: Statistical learning. *Dev. Sci.*, *14*(3), 464–473.

Baer, C., & Kidd, C. (2022). Learning with certainty in childhood. *Trends Cogn. Sci.*, *26*(10), 887–896.

Barr, D. J., Levy, R., Scheepers, C., & Tily, H. J. (2013). Random effects structure for confirmatory hypothesis testing: Keep it maximal. *J. Mem. Lang.*, *68*(3).

Bates, D., Kliegl, R., Vasishth, S., & Baayen, H. (2018). Parsimonious mixed models.

Bergstra, J., & Bengio, Y. (2012). *Random search for hyper-parameter optimization* (Vol. 13).

Blanco, N. J., & Sloutsky, V. M. (2020, September). Attentional mechanisms drive systematic exploration in young children. *Cognition*, *202*(104327), 104327.

Bullock, M. (1985). Causal reasoning and developmental change over the preschool years. *Hum. Dev.*, *28*(4), 169–191.

Bunge, S. A., & Wright, S. B. (2007). Neurodevelopmental changes in working memory and cognitive control. *Curr. Opin. Neurobiol.*, *17*(2), 243–250.

Charnov, E. L. (1976). Optimal foraging, the marginal value theorem. *Theor. Popul. Biol.*, *9*(2), 129–136.

Chevalier, N., Dauvier, B., & Blaye, A. (2018). From prioritizing objects to prioritizing cues: a developmental shift for cognitive control. *Dev. Sci.*, *21*(2).

Chevalier, N., Meaney, J. A., Traut, H. J., & Munakata, Y. (2020). Adaptiveness in proactive control engagement in children and adults. *Dev. Cogn. Neurosci.*, *46*(100870), 100870.

Cohen, A. O., Nussenbaum, K., Dorfman, H. M., Gershman, S. J., & Hartley, C. A. (2020). The rational use of causal inference to guide reinforcement learning strengthens with age. *NPJ Sci Learn*, *5*, 16.

Coughlin, C., Robins, R. W., & Ghetti, S. (2019). Development of episodic prospection: Factors underlying improvements in middle and late childhood. *Child Dev.*, *90*(4), 1109–1122.

Davidow, J. Y., Insel, C., & Somerville, L. H. (2018, August). Adolescent development of value-guided goal pursuit. *Trends Cogn. Sci.*, *22*(8), 725–736.

Decker, J. H., Otto, A. R., Daw, N. D., & Hartley, C. A. (2016). From creatures of habit to Goal-Directed learners: Tracking the developmental emergence of Model-Based reinforcement learning. *Psychol. Sci.*, *27*(6), 848–858.

de Eccher, M., Mundry, R., & Mani, N. (2024). Children's subjective uncertainty-driven sampling behaviour. *R. Soc. Open Sci.*, *11*(4), 231283.

DeMaster, D., Pathman, T., Lee, J. K., & Ghetti, S. (2014, November). Structural development of the hippocampus and episodic memory: developmental differences along the anterior/posterior axis. *Cereb. Cortex*, *24*(11), 3036–3045.

Deng, W. S., & Sloutsky, V. M. (2015, March). The development of categorization: effects of classification and inference training on category representation. *Dev. Psychol.*, *51*(3), 392–405.

Deng, W. S., & Sloutsky, V. M. (2016, December). Selective attention, diffused attention, and the development of categorization. *Cogn. Psychol.*, *91*, 24–62.

Fiser, J., & Aslin, R. N. (2002, November). Statistical learning of new visual feature combinations by infants. *Proc. Natl. Acad. Sci. U. S. A.*, *99*(24), 15822–15826.

Forest, T. A., Schlichting, M. L., Duncan, K. D., & Finn, A. S. (2023). Changes in statistical learning across development. *Nat. Rev. Psychol.*, *2*(4), 205–219.

Gershman, S. J., Blei, D. M., & Niv, Y. (2010). Context, learning, and extinction. *Psychol. Rev.*, *117*(1), 197–209.

Gopnik, A., Sobel, D. M., Schulz, L. E., & Glymour, C. (2001). Causal learning mechanisms in very young children: Two-, three-, and four-year-olds infer causal relations from patterns of variation and covariation. *Dev. Psychol.*, *37*(5), 620–629.

Griffiths, T. L., Navarro, D. J., & Sanborn, A. N. (2006). A more rational model of categorization. *Proceedings of the Annual Meeting of the Cognitive Science Society*, *28*(28).

Gualtieri, S., & Finn, A. S. (2022, April). The sweet spot: When children's developing abilities, brains, and knowledge make them better learners than adults. *Perspect. Psychol. Sci.*, 17456916211045971.

Harhen, N. C., & Bornstein, A. M. (2023). Overharvesting in human patch foraging reflects rational structure learning and adaptive planning. *Proc. Natl. Acad. Sci. U. S. A.*, *120*(13), e2216524120.

Hartley, C. A., Nussenbaum, K., & Cohen, A. O. (2021). Interactive development of adaptive learning and memory. *Annual Review of Developmental Psychology*, *3*, 59–85.

Jiang, N., Kulesza, A., Singh, S., & Lewis, R. L. (2015). The dependence of effective planning horizon on model accuracy. *Adapt Agent Multi-agent Syst*, 1181–1189.

Jung, Y., Walther, D. B., & Finn, A. S. (2021, September). Children automatically abstract categorical regularities during statistical learning. *Dev. Sci.*, *24*(5), e13072.

Karmiloff-Smith, A. (1995). *Beyond modularity.* The MIT Press, Massachusetts Institute of Technology.

Kemp, C., Perfors, A., & Tenenbaum, J. B. (2007, May). Learning overhypotheses with hierarchical bayesian models. *Dev. Sci.*, *10*(3), 307–321.

Keresztes, A., Ngo, C. T., Lindenberger, U., Werkle-Bergner, M., & Newcombe, N. S. (2018, August). Hippocampal maturation drives memory from generalization to specificity. *Trends Cogn. Sci.*, *22*(8), 676–686.

Lapidow, E., Killeen, I., & Walker, C. M. (2022, March). Learning to recognize uncertainty vs. recognizing uncertainty to learn: Confidence judgments and exploration decisions in preschoolers. *Dev. Sci.*, *25*(2), e13178.

Lapidow, E., & Walker, C. M. (2020). The search for invariance: Repeated positive testing serves the goals of causal learning. In *Language and concept acquisition from infancy through childhood* (pp. 197–219). Cham: Springer International Publishing.

Leslie, A. M. (1982). The perception of causality in infants. *Perception*, *11*(2), 173–186.

Liquin, E. G., & Gopnik, A. (2022). Children are more exploratory and learn more than adults in an approach-avoid task. *Cognition*, *218*, 104940.

Lucas, C. G., Bridgers, S., Griffiths, T. L., & Gopnik, A. (2014). When children are better (or at least more open-minded) learners than adults: developmental differences in learning the forms of causal relationships. *Cognition*, *131*(2), 284–299.

Mobbs, D., Trimmer, P. C., Blumstein, D. T., & Dayan, P. (2018, July). Foraging for foundations in decision neuroscience: insights from ethology. *Nat. Rev. Neurosci.*, *19*(7), 419–427.

Moreno-Bote, R., Ramírez-Ruiz, J., Drugowitsch, J., & Hayden, B. Y. (2020). Heuristics and optimal solutions to the breadth-depth dilemma. *Proc. Natl. Acad. Sci. U. S. A.*, *117*(33), 19799–19808.

Niebaum, J. C., Chevalier, N., Guild, R. M., & Munakata, Y. (2021). Developing adaptive control: Age-related differences in task choices and awareness of proactive and reactive control demands. *Cogn. Affect. Behav. Neurosci.*, *21*(3), 561–572.

Nussenbaum, K., & Hartley, C. A. (2019). Reinforcement learning across development: What insights can we draw from a decade of research? *Dev. Cogn. Neurosci.*, *40*, 100733.

Nussenbaum, K., Martin, R. E., Maulhardt, S., Yang, Y. J., Bizzell-Hatcher, G., Bhatt, N. S., ... Hartley, C. A. (2023). Novelty and uncertainty differentially drive exploration across development. *Elife*, *12*.

Nussenbaum, K., Scheuplein, M., Phaneuf, C. V., Evans, M. D., & Hartley, C. A. (2020). Moving developmental research online: Comparing in-lab and web-based studies of model-based reinforcement learning. *Collabra Psychol.*, *6*(1).

Perfors, A., & Tenenbaum, J. (2009). Learning to learn categories. *Proceedings of the Annual Meeting of the Cognitive Science Society*, *31*(31).

Potter, T. C. S., Bryce, N. V., & Hartley, C. A. (2017). Cognitive components underpinning the development of model-based learning. *Dev. Cogn. Neurosci.*, *25*, 272–280.

Raviv, L., & Arnon, I. (2018, July). The developmental trajectory of children's auditory and visual statistical learning abilities: modality-based differences in the effect of age. *Dev. Sci.*, *21*(4), e12593.

Renardy, M., Joslyn, L. R., Millar, J. A., & Kirschner, D. E. (2021). To sobol or not to sobol? the effects of sampling schemes in systems biology applications. *Math. Biosci.*, *337*, 108593.

Saffran, J. R., Aslin, R. N., & Newport, E. L. (1996). Statistical learning by 8-month-old infants. *Science*, *274*(5294), 1926–1928.

Schlichting, M. L., Guarino, K. F., Schapiro, A. C., Turk-Browne, N. B., & Preston, A. R. (2017). Hippocampal structure predicts statistical learning and associative inference abilities during development. *J. Cogn. Neurosci.*, *29*(1), 37–51.

Schlichting, M. L., & Preston, A. R. (2015, February). Memory integration: neural mechanisms and implications for behavior. *Curr. Opin. Behav. Sci.*, *1*, 1–8.

Schulz, E., & Gershman, S. J. (2019). The algorithmic architecture of exploration in the human brain. *Curr. Opin.*

Neurobiol., *55*, 7–14.

Schulz, E., Wu, C. M., Ruggeri, A., & Meder, B. (2019). Searching for rewards like a child means less generalization and more directed exploration. *Psychol. Sci.*, *30*(11), 1561–1572.

Shin, Y. S., & DuBrow, S. (2021, January). Structuring memory through inference-based event segmentation. *Top. Cogn. Sci.*, *13*(1), 106–127.

Shufaniya, A., & Arnon, I. (2018, November). Statistical learning is not age-invariant during childhood: Performance improves with age across modality. *Cogn. Sci.*, *42*(8), 3100–3115.